26th International System Dynamics Conference, Athens

Modeling as Theory-Building

Markus Schwaninger

Stefan Groesser

University of St. Gallen
Institute of Management
System Dynamics Group
Dufourstrasse 40a
CH - 9000 St. Gallen
Switzerland
Tel: +41 71 224 23 82 / Fax: +41 71 224 23 55
markus.schwaninger@unisg.ch
stefan.groesser@unisg.ch

July 2008

ABSTRACT

The purpose of this contribution is to make the idea of modeling as theory-building operational. We conceive the modeling process as a theory-building process, thereby opening up a new perspective on the methodology of modeling in the social sciences. By reconceptualizing the notion of modeling, we hope to convey the advantages of more conceptual thinking in management. Practitioners could gain effectiveness in dealing with external complexity if they would espouse the modeling task as a disciplined reflection and communication geared toward the elaboration of theories. This contribution is based on projects in which System Dynamics models for real-world issues were constructed together with corporate partners. To clarify the isomorphic nature of theory-building and formal modeling and illustrate the approach to modeling as theory-building, one of these modeling ventures is described in detail.

Keywords: Modeling and Simulation, Theory-Building, Model-based Management, Case Study

An earlier version of this article was published as:
Schwaninger, M. & Groesser, S. N. 2008. Model-Based Theory-Building with System Dynamics. Systems Research and Behavioral Science, in press.

I. INTRODUCTION

This paper explores the issue of formal modeling from a theory-building stance. Our specific goal is to make the concept of modeling as theory-building operational and demonstrate its great potential. This is based on many realized field experiments, one of which will be described here in a detailed case study. The question which guides our analysis is: "Can formal computer modeling be an effective approach to theory-building, and how?" When referring to computer modeling we take System Dynamics as an exemplary methodology.

System Dynamics (SD) modeling is often colloquially referred to as theory-building. Although theory-building in this context has been subject to very little discussion, we have identified some work in the area. Early on, Jay Forrester conceptualized System Dynamics modeling and simulation in essence as a discipline for theorizing, involving experimental work, "designed to prove or disprove the initial hypothesis" (1961: 450). Hanneman (1988) conceived computer-assisted dynamic modeling as theory-building. Schwaninger (2003) described the modeling endeavor as the building of local theories based on conceptual frameworks – the archetypes which emanate from the SD community (cf. Senge, 2006; Wolstenholme, 2003). Finally, Karlöf and Lövingsson (2006) equated practitioners' problem solving with theory-building. However, it is necessary to delve deeper into the concept of theory construction based on formal models. We will address this gap.

The paper is structured as follows. We will begin by conceptualizing theory-building, modeling and related concepts. Then we will explore a detailed case study, which refers to a modeling project as the locus of theory-building. A discussion will ensue, and the paper closes with brief conclusions.

II. CONCEPTS

Theory-Building

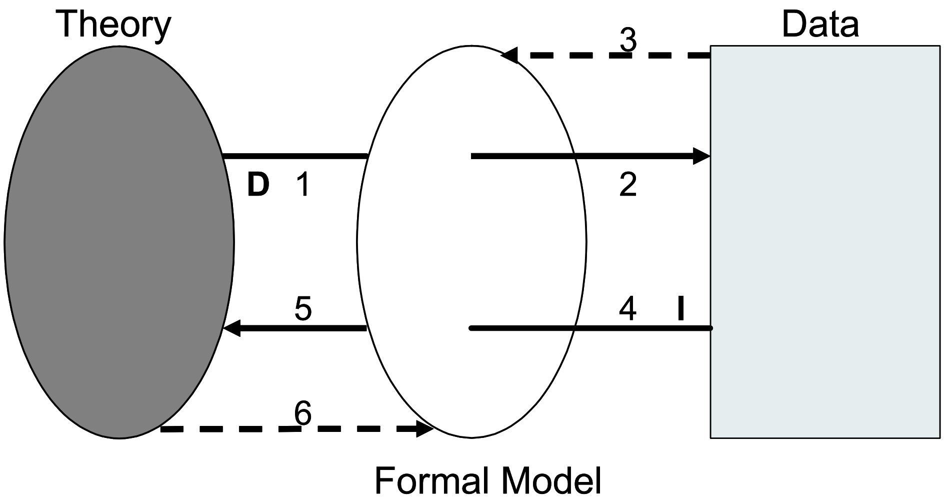
Theory-building, in principle, is more than an exercise in academic abstractions, but rather an activity fundamental to the survival of societies, organizations and even individuals (cf. Schwaninger and Hamann, 2005). We conceive a theory as a structured, explanatory, abstract, and coherent set of interconnected statements about a reality. These statements consist of constructs linked together by testable propositions that have a coherent logic and related assumptions. (Davis et al., 2007: 481; see also Miller and Page, 2007: 59). The term derives from the Greek 'theorein' – to look at, to ponder, to investigate. Concomitantly, the noun 'theoria' refers to notions such as observation, inquiry, reflection and research. Theorizing, in this sense, is to observe what is going on in the real world, to reflect or experiment about it and to draw systematic conclusions, which have practical implications.[1] Theory-building as understood here, consists of generating and formalizing a theory in order to orientate action.

However, there are different kinds of theory-building (see Figure 1). The differences become visible if we take the source of knowledge as the criterion of distinction. The common way of building theories is to rely on observation as the source of knowledge.

---

[1] Often a distinction is made between a reference to a system 'out there' and a subjective interpretation by a human (e.g., Checkland, 1981; Zagonel, 2004). The stance taken here is that the modelling of social systems as discussed in this paper always involves the subjective interpretation by an observer who, in this case, is a modeller.

One way is to gather the data by observing real-life events, studying cases, carrying out surveys, etc. The data are then analyzed, usually explored statistically, interpreted and presented. Successive tests then lead to the formulation of the theory. This is the *'data-to-theory'* process, which we call *induction*. The other approach is to build a theory, based on certain assumptions, using a logical sequence of steps. Then data is collected to test the theory. This is the *'theory-to-data'* approach which we call *deduction*. As this description shows, deduction and induction are components of the research process, which in principle complement each other. Figure 1 also shows that the construction (and the validation) of models has both, inductive and deductive components.



Code:

D: Deduction (1,2)
I: Induction (4,5)
- - - Validation (3,6)

1,4: Formation of formal model
2: Trigger for data search
5: Abstraction
3,6: Adjustment of formal model (i.e., change, alternatively falsification or corroboration)

Figure 1: Induction and deduction in modeling and theory-building.

When data are gathered the researcher relies on a theoretical framework, and when starting a deductive venture some real-life events have already been observed, which influence the choice of categories and the logic of the deduction.

As far as the theory-building processes we are interested in here are concerned, they do not involve deduction or induction alone, but are a combination of both methods. Methodologically, the processes are focused on building a formalized, quantitative model and conducting simulations with computer support. This is referred to as computer-supported theory-building. In the cases we will refer to, the System Dynamics methodology has been used as a conceptual and instrumental device for constructing models consistent with the principles of theory-building.

In principle, models are often theories or parts thereof.[2] Theories tend to have generic and properly formalized models as their constituent parts. Theories which emanate from a process of theorizing that is based on an explicit, formal model[3] have the potential to be stronger – in terms of both robustness and reach – than theories, which are largely based on implicit, mental models.[4] In this connection, Ashby's Law of Requisite Variety must be quoted: "Only variety can absorb variety".[5] In our context, this means that the strength of a theory hinges on its richness in relation to the complexity of the reality it deals with. In a similar vein, the Conant-Ashby-theorem indicates that the effectiveness of an operator cannot be higher than is the power of the model on which his or her operation is based.[6]

Models and Modeling

We conceive modeling as a process by which formal models are built. A formal model is an explicit representation[7] or a construction of a reality. If perception is an activity, as opposed to a passive happening (Neisser, 1976), then a model is a subjective construction. More precisely, it is normally a conceptual construction of an issue under study. Modeling, according to the constructivist position, is the construction of a subjective reality. The modeler is an observer who, by the act of observing or modeling, creates "a new world" (cf. von Foerster, 1984).

In this contribution we concentrate on computer models – System Dynamics models in particular. System Dynamics modeling is about constructing formal models of dynamic issues, as continuous feedback systems. They incorporate hypotheses about the causal connections of parameters and variables as functional units, and the outcomes of their interactions. If used for theory-building, SD models must be fully transparent, i.e., they must be available in the form of white boxes instead of black boxes, which would be counterproductive. The model becomes a strong device for supporting the process of theory-building, in which the model is, or gradually becomes, part of the theory itself. Hence, theory-building and model-building in the end are one, evolving in the form of a discourse, in which different people are involved. In other words, what keeps the discourse going is usually not a completed model, but a model in its different stages, built by a group of participants. Such group model-building has already been the object of deeper studies (Vennix, 1996; Richardson and Andersen, 1995).

---

[2] This expresses a view of theory-building in a wider sense, and is in line with Forrester's claim that a "general model" is "a theory of a particular class of system" (1968: 607).

[3] A formal model can be of quantitative nature (e.g. numeric), but it can also be of a qualitative type (e.g. graphical).

[4] The benefits of formal theory-building has been widely shown in the literature. See for example Lane, 2008; Hannan et al., 2007; Homer, 1996; Grove and Meehl, 1996.

[5] In the original: "Only variety can destroy variety" (Ashby, 1956).

[6] In the original: "Every good regulator of a system must be a model of that system." (Conant and Ashby, 1981). In this context it is noteworthy that, according to experimental research, the structures of mental models used by decision-makers are strongly predictive of their performance (cf. Ritchie-Dunham, 2002, and literature related therein).

[7] Such representation can be descriptive or anticipative or prescriptive.

The formalization of mental models by SD – qualitative and quantitative[8] – fosters transparency. It also increases their refutability, which according to Popper (2002) is a prerequisite for submitting them to scientific debate. If we follow Popper's critical-rationalist theory of science, then theory-building in the human and social sciences is not primarily meant to be an exercise of underpinning hypotheses, nor of 'proving' their truth. It is a process by which assumptions and theories, i.e., systems of hypotheses are specified and then submitted to tests. These tests are essentially endeavors of falsification. If falsified, then the theory (or hypothesis) is refuted. If, however, the attempt of falsification is not successful, then the theory can be temporarily maintained and grows in strength. There is a crucial difference between the falsificationist and the verificationist approach. The latter aims at creating absolute certainty, which is illusory. The former accepts that all theory is provisional, adhering to an evolutionary concept. While verificationism needs to accumulate corroborative evidence indefinitely, falsificationism is based on testing the robustness of theories by attempting to refute them.

The basic value of a model or a simulation outcome is that it embodies propositions that can be refuted. The point is not primarily whether a proposition is true or false, but that it provides an anchor around which arguments can be built.

System Dynamics models greatly enhance falsifiability – each interrelationship can be tested, both logically and empirically. In this sense, an SD model is, in principle, a candidate for a theory. This consideration is applicable to properly constructed models, i.e., those models which make the underlying assumptions explicit, operationalize the variables and parameters, and are submitted to adequate procedures of model validation (Forrester and Senge, 1980; Barlas, 1996; Sterman, 2000).

Not every model is a theory. Under what conditions does a model become a theory? A clear answer is given by Lane (2008): For a model to qualify as a theory, "what is required is a model along with a plausible account of why the model produces the behavior that it does."

Range of Theory

Academics are, in the first place, interested in creating *general theories*.[9] A general theory is any theory that attempts a highly generic, often overall explanation of a whole range of phenomena, e.g., of social systems. It often strives for a unification of several separate theories. An example from institutional economics would be a theory of social and economic change.

Middle-range theories are "theories that lie between the minor but necessary working hypotheses that evolve in abundance during day-to-day research and the all-inclusive systematic efforts to develop a unified theory that will explain all the observed uniformities of social behavior, social organization and social change" (Merton, 1968: 39). These middle-range theories consolidate different hypotheses or findings (Merton,

[8] We adhere to the view that also qualitative approaches to formalization do exist, e.g., graphical and verbal approaches, if the criteria for formalization are consistency, precision, inequivocality, etc.

[9] A similar concept used in the social and human sciences is the one of 'grand theory' (cf. Skinner, 1985).

1957: 328 and 280), and are empirically grounded. Continuing the institutional economics example used to illustrate general theory, an instance of a middle-range theory would address connections of specific hypotheses from both the economic and the social domains. In our view, the consolidation provided by a middle-range theory could also be one of different local theories, e.g., by expanding the realm in which a theory applies (see also Paich, 1985).

Local theories are those theories that apply to a particular context, explaining behaviors encountered in specific instances. An example here would apply the principles of economic change to a specific social system, e.g., an organization or society, by explaining how the current situation of that unit emerged. In this sense, theory-building is not only the domain of theoreticians, but also of practitioners.

In our case example the focus will be on middle-range theories. Most other parts of the text are also applicable to general and local theories.

Criteria for Evaluation

The criteria for evaluating a theory-building process should, in our view, refer to the outcome of that process, i.e., the quality of the resultant theory. A pertinent set of eight criteria for theory evaluation was developed by Patterson (1986). The original list has been completed with definitions by Holton and Lowe (2007). These are included here with small modifications and extensions to the set of criteria pertinent to model-based theory-building (criteria one and eight have been added):

1. Refutability: ability of a theory to be falsified (refuted) or supported;
2. Importance: a quality or aspect of having great worth or significance; acceptance by competent professionals may be indicative of importance;
3. Precision and clarity: a state of being clear; hypotheses can easily be developed from the theory;
4. Parsimony and simplicity: uncomplicated; limitation of complexity and assumptions to essentials;
5. Comprehensiveness: covering completely or broadly the substantive areas of interest;
6. Operationality: specific enough to be testable and measurable;
7. Validity: valid, accurate representation of the real system under study;
8. Reliability: free of measurement errors;
9. Fruitfulness: statements are made that are insightful, leading to the development of new knowledge;
10. Practicality: provides a conceptual framework for practice.

In the discussion of the case study, we will revert to this list for the evaluation. We have examined the ten criteria and found that they are not only pertinent for the evaluation of a theory. Our conclusion is that this whole set of criteria is equally appropriate for the evaluation of any formal model, even if it does not have the status of a theory.

III. CASE STUDY: BUILDING A THEORY OF PRODUCT LAUNCH STRATEGIES

Context

The following case study analyses a modeling project in which a System Dynamics simulation model for a real-life management challenge was developed. The model will be evaluated from the perspective of theory-building utilizing the evaluation criteria laid out in the previous chapter. It stems from 36 case studies, which were essentially theory-building endeavors. In each one, a local or a middle-range theory was developed. The modelers were students who were engaged in SD courses at the University of St. Gallen, Switzerland. The modeling ventures were undertaken in collaboration with external partners (industrial or service firms, including consultancies, public organizations, and a local community). They were supervised and coached by of our team; one faculty member (M.S.) and one or two assistants (S.G. or others). The ventures were accomplished within an eleven-year period. We present a revelatory[10] exemplar, i.e., an instance, which illustrates something that is – in principle – encountered in all of our cases: the specific nature and features of a theory-building process, in particular, the combination of deduction and induction, while complying with the criteria of high quality theory and good modeling practice.

The project under study was carried out in cooperation with a Swiss industrial company, which is one of the two world leaders in its field. The company faced a number of problems at the time of a product launch: it had no means of estimating the potential sales development; it lacked knowledge about the factors that significantly influence the development of product sales; and it was not able to assess the possible impact of its management activities. The project was highly relevant because the short product life cycles require fast and resolute decisions. The purpose of the model was to support the company's management of product launches in order to optimize the return on investment of each product generation. The aim was to understand which factors influence the product launches; in particular, to identify those that have the strongest impact on the financial outcome of a product launch. Moreover, the model should yield general policy insights to guide future product launches. For us, the external partners, the crucial objective of the project was beyond mere practical empiricism; the search for a theory of industrial product launches pervaded the whole endeavor.

The project started in May 2007 and lasted for six months. The interaction with the company hinged on two gatekeepers who were the main interlocutors of the external modelers, i.e., the authors with the support of three students, for whom the modeling was part of their training. The modeling work was carried out by the external partners in collaboration with internal gatekeepers, who then involved additional company members. For the sake of brevity and clear distinction, we will use the terms *core modeling team* and *researchers* for the external group, while the internal people involved actively in the modeling process will be called *client team*.

---

[10] The use of single case studies, as opposed to multiple case studies, is indicated if they are revelatory (Yin, 2003: 45f.).

Modeling Process

The modeling process followed standard procedures as laid out, for instance, by Richardson and Pugh (1981), Wolstenholme (1990), and Sterman (2000). The core modeling team interacted with the client team in several Group Model Building sessions. Before the first session, an issue statement as well as a purpose statement of the model had been crafted in an internet dialogue between the internal gatekeepers and the external partners. These preliminary statements were discussed, changed and constituted in the first joint session. The group effort, on that occasion, led to a first conceptualization of the dynamic hypotheses, which will be expounded shortly.

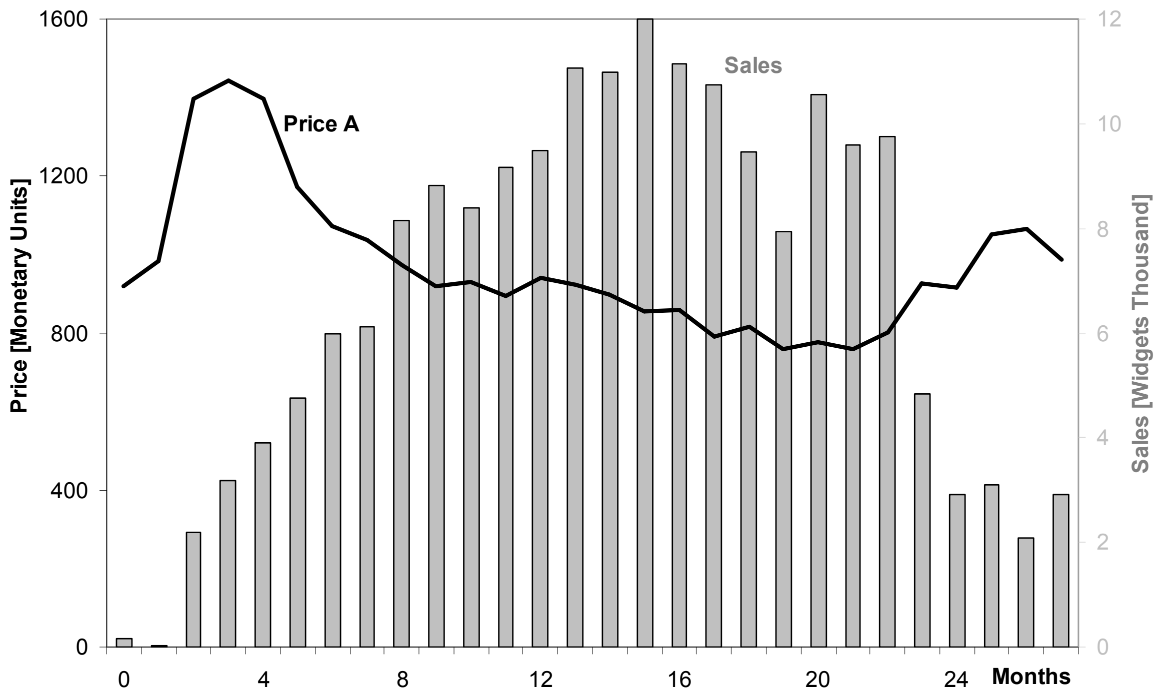


Figure 2: Data about product sales, used as a reference mode for the product launch model, in relation to the exogenous price of the product.

In addition, reference modes for two variables – monthly product sales rate and revenues – were established to guide the modeling endeavor. Figure 2 shows the time series for sales rate based on the empirical values supplied by the company. In the following sessions, the qualitative dynamic hypotheses were formalized as quantitative, equation-based dynamic hypotheses, which were subject to scientific discourse, falsification, rejection and reformulation. The clients were 'walked through' the model in detail in order to ensure both the model's face validity and the acceptance by our corporate partners, – essentially a deductive exercise. These were 'theoretical' sessions for practical probation. The reflections made on these occasions led to improvements of the early model versions and to the identification of additional data requirements. Several subsequent modeling sessions helped to set parameters and test the model, as well as perform scenario and policy analyses. Interactions with the client team involved personal and online collaborations as well as offline correspondence.

In the early stages of the modeling endeavor process facilitation and communication with the client was given priority over the precision of the model. It was particularly important that the parameters used were reasonable to the problem owners. In later stages, after the client had become more familiar with the modeling process and technique, we elicited more accurate data, required for the parameterization of the model, in additional online meetings. During the whole modeling process, the modelers focused on keeping the model as parsimonious as possible while avoiding black box modeling, which could have easily occurred given the complexity of the market situation and the company's size, structure and product portfolio. Figure 3 shows the main feedback structure of the model that emerged within about four months. The diagram also highlights that the core of the model is a Bass diffusion structure (cf. Sterman, 2000).

To avoid overloading this contribution, we refrain from presenting the whole model in all its details, which are definitely much richer than what can be shown in Figures 3 and 4. In this context we will not present all of the substantive theories about product launches that were used in support of the model building exercise. The main sources were Rogers (2003) and Sterman (2000) concerning product diffusion, Warren (2002) with respect to the structure of product launch processes, Zirger and Maidique (1990) as well as Henard and Szymanski (2001) in relation to success factors in product launch.
The purpose of the initial modeling was to develop an explanation of the historical sales of a typical product of the company, as graphed in Figure 2, and not to investigate the future. We started out with initial hypotheses, which were rejected and replaced by new ones, and maintained or eliminated in an evolutionary process of iterative, scientific modeling, in which some of the hypotheses 'survived'. The hypotheses we present here passed numerous trials of falsification with respect to the empirical data. One of those dynamic hypotheses was that growth in the number of customers, and thus growth in the monthly sales rate, was the result of improvements in the perceived attractiveness of the company's product relative to the products of competitors, rather than reputation effects (Equation 1).

$$\mathrm{AR_t} = \max\left\{ \mathrm{ARA_t} + \mathrm{AWM_t}, \mathrm{PC_t} \right\} \quad (1)$$

with the adoption rate, AR_t, being the maximum between the adoption from relative product attractiveness, ARA_t, plus the adoption from word of mouth, AWM_t, and the potential customers, PC_t.

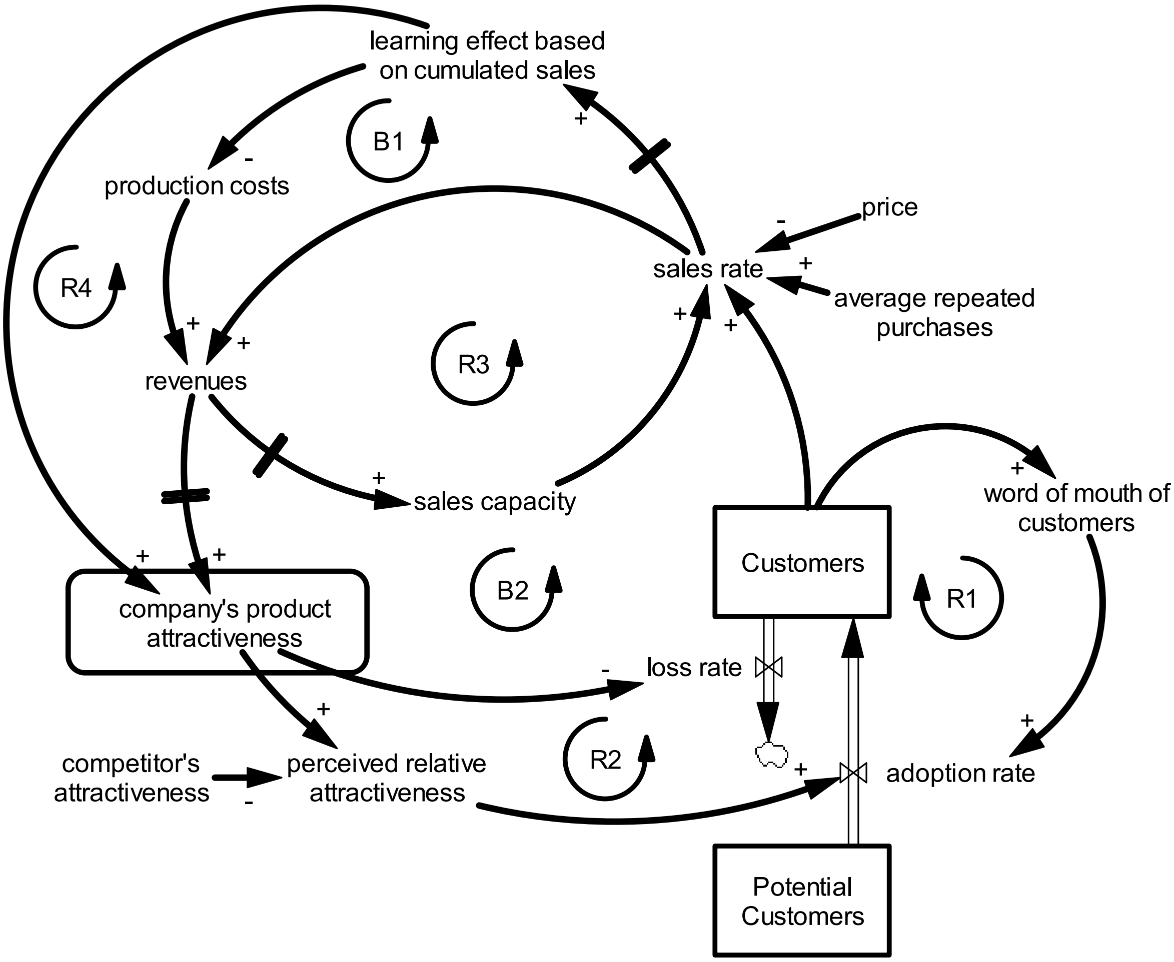


Figure 3: Main feedback structure (condensed) of the product launch model. The rounded rectangle (left side) summarizes several product attributes that are detailed in Figure 4.

Attractiveness was defined as a function of quality of service, product features, effect of marketing activities, and product assembly time, all of which were funded or influenced by the company's revenues. The influences of all of these attributes on the product attractiveness were weighed according to the judgment of the company's product experts. The product assembly time was positively influenced by cumulative learning effects; Figure 4 provides details about the attributes of the company's product attractiveness, by means of a more detailed representation than in Figure 3 of the loop 'Company's product attractiveness → Customers → Revenues → Company's product attractiveness'.

A second dynamic hypothesis was that only a positive influence of word of mouth on the adoption rate existed (R1, Figure 4); we thereby abstracted from possible negative effects of word of mouth (Equation 2).

$$AWM_t = \frac{\bar{c} * \bar{i} * (TP_t - PC_t) * PC_t}{TC_t} \quad (2)$$

with adoption from word of mouth, AWM_t, calculated by the contact rate, c, the adoption fraction, i, the total customers, TC_t, and potential customers, PC_t.

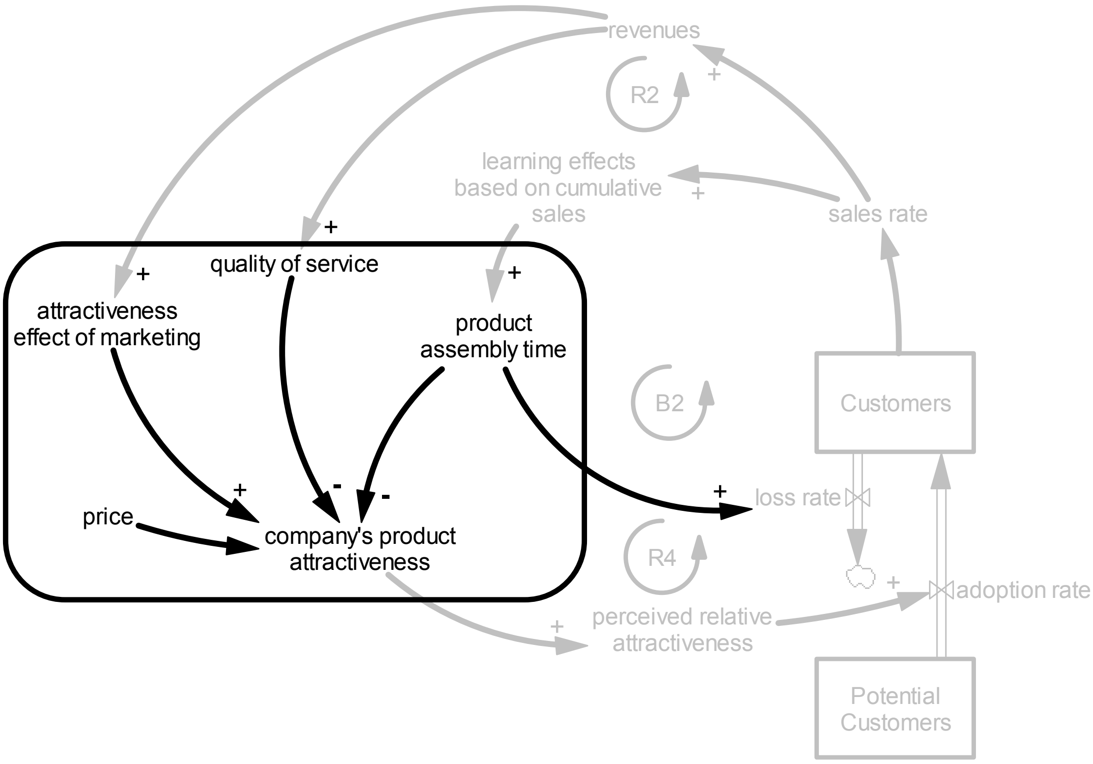


Figure 4: Detailed representation of the attributes of the company's product attractiveness (for the complete feedback loops see Figure 3).

The negative impact of long product assembly times, leading to reduced product attractiveness and loss of customers, was also included in the model (see B2, Figure 4). The last dynamic hypothesis was that the sales volume depended on the sales capacity, which is a function of the company's revenues (see R3, Figure 3). Other determinants of sales volume were price and average repeated purchases (Equation 3).

$$SV_t = \min\left\{ FTP_t + RP_t, SC_t \right\} \quad (3)$$

with sales volume, SV_t, the value of first time purchase; FTP_t, the value of repeat purchases, RP_t, and sales capacity, SC_t.

All three dynamic hypotheses stood the empirical tests, as will be delineated in the next section. During the whole modeling process, the modelers sought to comply with the aforementioned scientific criteria for theory-building. We elaborate on this in the next section.

Model Validation
As suggested by System Dynamics experts (e.g., Barlas, 1996; Forrester and Senge, 1980; Sterman, 2000), the validation of the model was an integral part of every modeling activity. During the iterative model development, we constantly questioned our hypotheses by means of direct and indirect structure tests as well as behavior-oriented tests. Each attempt to falsify the provisional system of hypotheses resulted in a better understanding of the model structure and the relationship between structure and behavior in the specific instance. These trials also helped the core modeling team to refine and calibrate the model, whereby confidence in the resulting simulations and structural analyses was enhanced. We will concentrate on four validation issues.

The first, and perhaps most important, aspect of validation was that the model's structure and its behavior emerged directly from interaction with the group of managers who would be using the model to support their decisions. The shared understanding of how a product launch occurred and what effects it would have on the company and potential customers, emerged from discussions. Initial sketches were drawn by hand on a whiteboard or on a projected computer screen. This helped establishing the credibility of the simulation model and nurture a feeling of ownership of the model among the group (see also: Richardson and Pugh, 1981: 355; Vennix, 1996). The inclusion of the client team in the modeling phase is a procedural measure to improve the validity of the model – it is not a single test of validity, as superficial reading of pertinent sources, e.g., Barlas (1996), Forrester and Senge (1980), might suggest.

Secondly, the inductive validation of propositions and hypotheses about causal relationships and assumptions about parameter values utilized the standard direct structure tests and drew on a variety of data sources. Following Randers, "model testing should draw upon all sources of available knowledge" (1980: 295), we used numerical data from the company's records. However, several parameters and functional relationships were not directly available in the existing data records. We had to rely on additional sources: (1) specific observations of company processes by members of the client team, (2) expert interviews, and (3) surveys of experts who were non-core project participants. It is noteworthy that the need for the additional data emerged during the model creation process. We sampled the data as required in order to validate new assumptions or causal relationships. The gatekeepers supported the modelers in this process and gathered independent estimates from knowledgeable managers. These data were analyzed for consistency and reliability before they were utilized for the model formulation. This way virtually all data needs were met and the researchers were able to create a simulation model on solid empirical grounds that is acceptable to the client.

Thirdly, sensitivity analyses and other indirect structure tests (also called structure-oriented behavior tests; Barlas, 1996) enhanced the confidence of managers in the simulation model. The analyses supported the validation of model assumptions, for which it was difficult to obtain numerical empirical data, by demonstrating the model's behavior insensitivity to changes in parameters. Moreover, the analyses showed that the derived policy insights were largely independent within plausible ranges of parameter values. One of these insights was that the product launch strategy of the competitors – either parallel or sequential – severely impacts on the company's sales development over the life cycle. The reason is that an early mover advantage can hardly be compensated for over the relatively short life cycle of 24 months.

Finally, behavior reproduction tests compared the reference modes that were based on historical real-world company data with the output of the simulation model. Forrester and Senge suggest a battery of tests in this category. However, most of them are not applicable to models that try to replicate one instance of life cycle dynamics, as in our case (Forrester and Senge, 1980: 219). Hence, we applied the behavior reproduction test in its basic version and compared the reference mode pattern with the pattern of the model output. The model reproduced the product life cycle pattern with high accuracy. Figure 5 shows the historical and the simulated time series for the product revenues (based on n= 26 data points, R^2=0.9967 for the corrected historical time series). To assess the model's behavior validity, we utilized the mean square error (MSE) and Theil’s inequality statistics (Sterman, 1984). MSE provides a measure of the total error and Theil’s statistics specify how it breaks down into three components: bias (U_M), unequal variation (U_S) and unequal covariation (U_C). An inspection of the historical data series reveals an interesting deviation from the normal life cycle behavior for $t \in [20, 22]$. Discussions with company experts could not clarify the reasons for the exceptional behavior. Hence, we treat the data points as outliers and substitute them with the average value of the long-term trend. Given the corrected historical data series of revenues, the MSE is 0.35%, the inequality statistics are: U_M = 0.01, U_S=0.01, and U_C=0.98. The major part of the error is concentrated in the unequal covariation component, while U_M and U_S are small. This signifies that the point-by-point values of the simulated and the historical data do not match even though the model captures the dominant trend and the average values in the historical data. Such a situation might indicate that the majority of the error is unsystematic with respect to the purpose of the model, and it should therefore not be rejected for failing to match the noise component of the data (Sterman, 1984: 56). The residuals of the historic and simulated time series show no significant trend and strengthen the assessment that the structure of the model captures the fundamental dynamics of the issue under study.

The final version of the model is relatively small, containing about 50 equations. However, it not only provides a close fit to the available product-level data, it does so for the right reasons: a set of parameter values and causal relationships that appeared robust and accurate to the client. Even though the model is small and condensed, it has triggered important insights. That was what had been required at the outset.

The model provides a theory that explains the evolution of the sales of a typical industrial product launch as a function of a set of drivers, i.e., parameters, which are the explicit expression of certain assumptions. Alternative hypotheses were proposed along the way, but were successively falsified and replaced by better ones. The selection filter which orientated the choices made en route was the process by which the internal validity was ascertained. Aspects of external validity and generalizability will be treated in the following section.

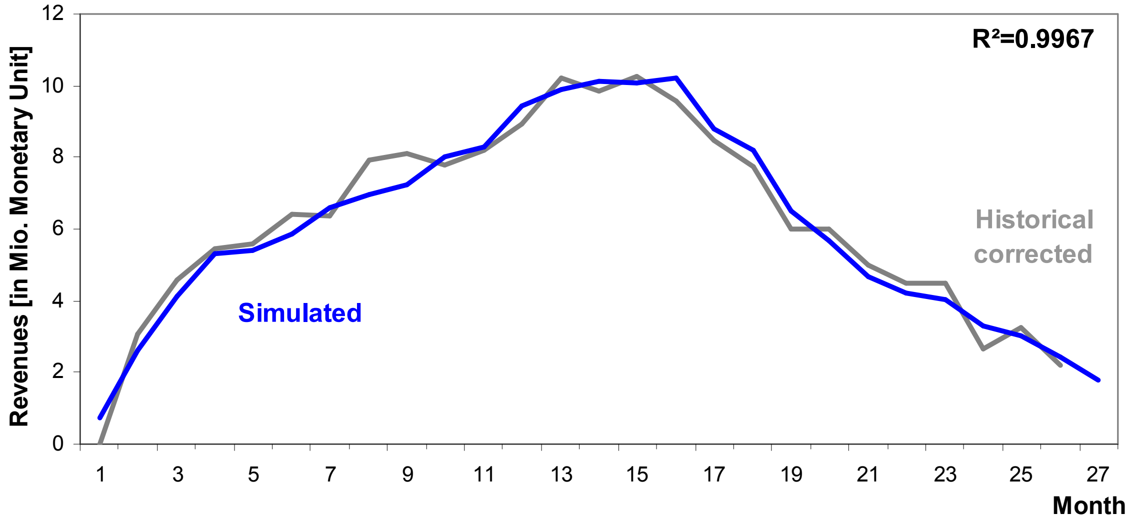


Figure 5: Comparison of the historical and the simulated data. The model explains approximately 99.7% of the behavior of the time series.

IV. DISCUSSION AND EVALUATION OF THE CASE

Our aim for this section is to discuss the modeling process and its results in terms of the conceptual-theoretical considerations laid out in the first part of the paper. We chose a single case setting, selecting a revelatory case from a sample of 36 cases. Our main argument is that System Dynamics modeling can result in a well crafted theory when the modeling process adheres to rigorous methodological criteria. We propose neither that this applies to SD models only nor that every model is per se a theory, let alone a theory of high quality. Davis et al. (2007) have shown that other methodologies also possess the potential for the construction of model-based theories. We have only supported our argument by means of the evidence collected in multiple applications of System Dynamics.

The purpose of our work was to build a theory applicable not only to a specific product but to any product launched by the case company. Beyond that, we tried to identify the class of systems in terms of situations to which the model could be applied (cf. Bell and Senge, 1980: 66; Lane and Smart, 1996: 94), and to model that particular class. We are confident that we have produced a highly generic system structure that can be used to understand not only the case described here, but virtually any industrial product launch. To test this assertion, we used a second set of reference data representing a product launch in a geographically distant country. The calibration process required only the adaptation of parameter values of the Bass diffusion structure (contact rate, adoption fraction, initial potential customers) and some product attributes, such as, repeat purchase rate, initial purchase rate, and price of product. This is plausible since the two countries (one within Europe, the other within North America) certainly differ in their information dissemination characteristics. This second application follows what Yin calls "replication logic", which allows cumulating knowledge across cases and building support for the theory under test (Yin, 2003: 37). Therefore, the results of the second application have increased our confidence that the created theory is applicable to a whole class of cases – product launches in industrial firms. In this sense, it is more than a local theory. As it consolidates several hypotheses and applies to a large range of cases, it is closer to a middle-range theory.

In the following, the quality of the attained theory will be evaluated utilizing the ten evaluation criteria of the quality framework developed in the first chapter:

- *Refutability*. A *sine qua non* of quantitative System Dynamics modeling is the formulation of an explicit set of mathematical equations. These are testable hypotheses, which enable researchers to check how well their assumptions match available data about the overall system behavior. In the simulation model, the underlying structural and behavioral assumptions are formalized and therefore testable. In other words, they can in principle be falsified.
- *Importance.* This is an attribute of a theory that can only be assigned by the target group for which the theory has been created. In our case, the primary targets were the project managers of the Swiss production company. They continually alleged during the modeling sessions that from their perspective the theory under construction was very helpful in addressing an important managerial challenge. This demonstrates the acceptance of the theory by the client organization. Towards the end of the year, after the formal end of the cooperation, the project was presented during a strategy meeting to the top management. The gatekeepers assured us that the next iteration of the project would, most likely, be supported by the top management executives. The importance of the project was evaluated as rather high by the gatekeepers.
- *Precision and Clarity*. Throughout the whole project, the modelers paid close attention to the consistent and precise use of the terms and concepts that appeared in the emerging theory. Moreover, the specific concepts were operationalized, explained and documented in a glossary to increase the theory's clarity. In the end, the theory furnished a precise and consistent conceptual apparatus. Furthermore, the relationships between the theory's concepts are defined explicitly and unambiguously.
- *Parsimony and Simplicity.* This criterion, in essence, advocates the avoidance of complication as opposed to the capability to absorb complexity, which has to be fostered as Ashby's Law postulates (see above). Theories therefore should be as simple as possible, but not simpler than necessary. In principle, System Dynamics is a methodology to capture dynamic complexity, not the complications of combinatorial complexity (see also Sterman, 2000). One of the dangers of SD modeling, however, is to succumb to pressures to create overly large models that try to contain the combinatorial richness of the real system, but not the dynamic complexity. The guideline to model a problem or issue, not a system with all its details (cf. Richardson and Pugh, 1981; Sterman, 2000), aims precisely at avoiding that trap. In the case under study, the modelers abided by this principle, and they had to defend it against severe initial resistance from the client team. The team members wished to incorporate more and more details, which conformed to their day-by-day experience and to an event-oriented perception of the situation, into the model. Finally however, the parsimonious approach proved to be appropriate and useful. The parsimony of the theory is expressed in the limited size of the model. Its simplicity lies in the small set of components, which are fully transparent and focused on the essentials.

- *Comprehensiveness.* This criterion examines whether the theory is sufficiently broad to cover the substantive issues of interest. In our case, we are confident that the comprehensiveness of the theory is reasonable because the boundary of the simulation model was repeatedly tested to see whether the scope of the model was adequate. At the beginning of the project the scope was rather narrow. During the course of the project, more details were added until the model contained structural elements the deletion of which hardly influenced the behavioral outcome. At that stage, the modelers began to regularly apply the boundary adequacy test (Barlas, 1996) in order to examine the appropriateness of the level of detail of the model as well as its scope. For instance, in an early version of the model the sales capacity had been modeled in more detail, but it became clear that the dynamics of hiring-firing and training of the sales force was not important to the situation-in-focus. We discovered that it was far more important to model the physical capacity of the product assembly, since this was one of the limiting factors for product customization and delivery: production capacity severely influenced the perceived relative attractiveness of the company's products. In conclusion, through several adaptations of the model boundary, which were based on empirical data and discussions with the client team, we were able to create a rich theory. It captured the essential dynamics of the situation under study, which indicates a moderate degree of comprehensiveness.
- *Operationality.* Operationality signifies that the theory is specific enough to be both testable and measureable. First of all, we made specific efforts to abide by the imperative that no variables should be used that do not have a counterpart in the real world. For example, an initially used variable 'costs ratio' was discarded, because it was not easily observable in the real world and it was actually not a variable used by the decision makers. We substituted total unit costs for costs ratio. Since the product launch theory is one of a socio-technical system, it incorporates not only clear-cut factors, but also soft variables that are difficult to measure (see Stouffer et al., 1950). For the development of the theory, we captured soft variables by using Likert scales (Cowles, 2001). For example, a questionnaire was distributed to knowledgeable experts in the client company. Their task was to quantify the impacts which different product attributes have on product attractiveness. The created glossary testifies that each variable of the product launch theory is highly operational as it can be measured and tested using real-world data.
- *Validity.* Validity implies that the model depicts the real system under study with a high level of accuracy. In System Dynamics, valuable research about model validation has been conducted (cf. Barlas, 1996; Forrester and Senge, 1980; Sterman, 2000). The extensive validation work undertaken for our theory of product launch included direct and indirect structure tests. In this case, we conducted an iterative process of model construction. The assumptions were validated or falsified in many simulation experiments based on comparisons between the model output and available data, especially the reference modes about monthly sales rate (cf. Figure 2) and revenues. For instance, the cost/sales structure of the initial model only contained direct sales and cost of sales for the tangible products. The validation against the reference modes indicated that a second source of revenues existed, which increased in its volume over time. In discussions with the client team, we identified product related services as this second

source. The dynamic change of the revenues was tied to increasing marginal revenues that were traced back to learning effects. The inclusion of this source of sales improved the accuracy of the model output compared to the reference modes. To summarize, based on numerical, written and verbal data, dynamic hypotheses were inductively created, implemented in the simulation model, and then deductively tested and compared to the reference modes. Innumerable trial and error experiments were necessary to arrive at a model structure that appeared to be robust and was able to reproduce the behavior of interest both accurately and for the right reasons. Homer (1996) refers to such iterative processes of induction and deduction as *scientific modeling*. The external validity of the model was examined as outlined at the beginning of this section. The results speak in favor of the generalizability of our findings.

- *Reliability.* In the context of theory evaluation, reliability can be understood as the extent to which the consequences of a test of the theory remain constant, if repeated under identical conditions. The product launch theory, for example, is a system of deterministic difference equations. It follows that reliability is warranted since re-tests of the theory will yield identical outcomes. However, reliability can also relate to the process of theory-building and its replication. A major provision for increasing the reliability of case studies is to thoroughly document the process of conducting research (Yin, 2003: 38 and 105), in our case, the development of the simulation model. We documented our simulation model in such a way that an external observer is able to trace the model development from the conceptualization phase until the final version.
- *Fruitfulness.* The question here is, whether the modeling and theory-building leads to important insights and ultimately to the creation of new knowledge. We are looking at the fruitfulness of a theory in terms of its heuristic power, as defined earlier in this paper. In order to increase the heuristic power of the theory, we tried to broaden the base from which the interpretation and conceptualization work was done: the independent researchers and the members of the client team held several Group Model Building sessions, in which the conceptualization and formulation of the model was conducted. Throughout the modeling process, and especially towards the official end of the project, the model was used extensively to explore what-if situations in order to draw inferences and explore management implications. These were discussed in the group setting where several new insights emerged. One of them concerned the role of product assembly time. This had been assumed to have an insignificant effect on company sales. The model gave a different answer. Product assembly time was the most important factor, the impact of which even surpassed product attractiveness. An examination of company data corroborated this counterintuitive result of our simulations. A second insight, as previously mentioned, concerned the severe impact of competitors' modes of product launch on the development of company sales. These insights provided new and important knowledge for the client. Finally and perhaps most importantly, our model integrated several partial theoretical components (from production theory and marketing) concerning product launches, producing a more comprehensive understanding of the subject under study (see also above: *comprehensiveness*).
- *Practicality*. Practicality means that a theory provides a conceptual framework for practice. The project team in our case consisted of researchers and responsi-

ble managers of the client company. Thus, a prerequisite for a theory under development to be of practical relevance was established. Furthermore, the members of the client company were included in each modeling activity. By this, it was ensured that the model incorporated concepts that both explain the dynamic behavior of interest and can, to a large degree, be influenced or manipulated by the decision makers. The process generated a theory that provided decision makers with a practically relevant framework of the essential control and policy variables.

In the Appendix we provide a table, which specifies, for each one of these criteria, the measures – that can be taken to achieve a high-quality theory. We refer to both modeling principles (Sterman, 2000, Vennix, 1996) and validation procedures (Schwaninger & Groesser, 2008).

It is not possible to fulfill all of the ten criteria to the maximal possible degree since some of them exhibit tradeoffs; for instance, *parsimony/simplicity* and *comprehensiveness*. While the first favors theories with as few assumptions as possible, the second tries to widen the scope of the theory to include different subject areas and more detail. In the case under study, a good balance between the degree of achievement of each criterion individually and in combination has in our view been attained.

The elaborated theory is more general than it might appear at a first glance. It refers to more than expected at the beginning. At the outset, the purpose of the project was to build a model that would support short-term decision making about product launches. The outcome provides decision makers with more than that: a theory about the interrelationships between the factors essential to a class of product launch situations and their longer-term implications. This is closer to a theory of the middle range than to a local theory.

The model-building procedure, as described in the case study, was one of scientific modeling, which "is distinguished from other approaches largely by the quality of evaluation and revision performed and by an insistence upon empirical evidence to support hypotheses and formulations" (Homer, 1996: 1). The final model falls into the category of 'canonical situation models', being fully formulated and calibrated. It is a case study reduced to its essentials, so as to enable the causal explanation or theory of the dynamic behaviors generated by the underlying structure (Lane, 1998: 937). We may assume that it is applicable to a whole class of systems, but we do not suggest that it proposes universal laws (cf. Lane, 2000). The mature model can be considered a theory[11], whereas the model in its pre-mature stages can not. It is an integrative theory, which builds on components, e.g., the Bass Diffusion Model and the Learning Curve.

[11] With the attribute *mature* we denote a model that is extensively validated and which largely fulfils the criteria for the assessment of theories as presented above.

V. CONCLUSIONS

Our aim was to make the theory-building approach to modeling operational.

The case studied here in detail as well as many of the other 35 cases we have referred to are by and large successful modeling ventures carrying the features of theory-building. They generated concrete benefits for real organizations. The outcome of these modeling processes is often a theory, i.e., a structured, explanatory, abstract, and coherent set of statements about a partial reality. In the modeling ventures drawn upon here, System Dynamics was the methodology used. All cases, the detailed one in particular, draw on a concept of theory which is evolutionary. The emphasis is on theory-building, not on static theorizing. Here, theorizing is about an interaction between modelers and model – a dialogue through which the theory is created. This view of modeling is not completely new. However, there is some innovative content in the idea of modeling as theory-building: It can lead to modelers gaining a wider understanding of the role and the methodology of model validation. There is a marked difference between exercising validation as a mere procedure by which a set of tests is carried out, and practicing validation as a reflective process by which a theory is built along epistemological lines.

Finally, we are getting back to the research question posed at the outset. Our answer is affirmative: Yes, formal computer modeling can be an effective approach to theory-building. Effectiveness here means that the formal modeling fosters the formation of high-quality theories. After our analysis we also dare to answer the "How"-addendum in the question. A necessary and perhaps even sufficient condition is that the modeler/theory-builder abide by the principles of good modeling practice, particularly in model validation (cf. Sterman, 2000).

We cannot present a comprehensive evaluation of the dynamic modeling approach in comparison with alternative approaches. This would have been beyond the scope of the study.

In the following, we summarize our insights and formulate a set of recommendations. These do not constitute a theory, but heuristic principles for modeling as theory-building, from the perspective of System Dynamics. We condense them in seven assertions which – in the light of the experiences over an eleven-year period – appear to be very important:

1. Modeling as theory-building: Good modeling and theory-building are essentially isomorphic in nature. They combine deduction and induction as methodological strategies. They are falsificationist rather than verificationist and aim at general insights. In addition, they are formalized to ensure continuity and improvements. The quality, namely the validity and the usefulness for achieving the model purpose, can be substantially increased if models are built in accordance with theory-building principles. To sum up, we plead for a theory-minded approach to model-building.

2. Formalization: The principle of formalization should be increasingly applied to model-building. Formal models enable appropriate testing of emerging insights. Their use also increases the likelihood that progress is made in the theory-building process, because systematic testing enhances improvements via selection. In contrast, refraining from formalization entails the danger of logical flaws and lack of precision.

3. Generalization: Generalization is a relevant principle of theory-building even if one does not aim at a general theory. Most models should strive for insights which are generic, i.e., which can be transferred to other situations of the same type. This way, external validity is enhanced, which also means that the model attains higher value. Be as general as possible and only as specific as necessary.

4. Validation: The quality, and with it the heuristic power of a model, is a function of its validity. Validity is the degree to which a model represents what it is supposed to represent. In other words, it is the degree to which the operational definitions of the variables and functions of the model accurately reflect the variables and functions they are designed to represent.[12] Be meticulous about model quality. Strive for improving the validity of the model continually in order to enhance the correspondence between the evolving theory and the reality it is supposed to explain (or "create"[13]). A well developed set of model validation tests to support theory-building is already available as part of the SD methodology.

5. Explanation: For modeling in general and theory-building in particular, the reproduction of a behavior under study is not enough. Beware of mechanical validation. Make sure that the model does not only reproduce, but that it enables the explanation of the behavior under study. For a model to qualify as a theory, it must include a plausible explanation of why the behaviors it produces are as they are (Lane, 2008, in press).

6. Falsification: The touchstone of a theory is falsification. The status of a theory can only be claimed, if serious attempts to falsify it have been undertaken and have foundered. To be able to meet that criterion, a theory must be formulated in a way that it is falsifiable. From that point of view – the critical-rationalist stance – the correct way to build a theory is by attempts at falsification, not by the justification of hypotheses. The fundamental importance of this very principle can be generalized for model-building of all kinds. Formal models are probably the most powerful device for systematic trials to falsify a theory and – from a theory-building perspective – that is their primary advantage. If modelers want to adhere to falsificationism rather than verificationism, thereby avoiding the confirmation trap, they will have to readjust their mindsets. In compensation they can fully utilize the potential of System Dynamics for building high-quality models.

7. Process Design: The design of the theory-building process is crucial for the quality of the resulting model. The trial of falsification should be iterative, triggering the selection of hypotheses and theories. As has been emphasized in System Dynamics research for many years, Group Model Building is a key to this design process. Internal and external experts have to cooperate. Experts with different disciplinary backgrounds should synergize.

[12] A validity of 100% is, in principle, unachievable for models of social systems (cf. Sterman, 2000).

[13] Models do not only represent what is, but they can also play a significant role in the construction of a new reality (cf. the constructivist position, e.g., von Foerster, 1984).

The design principles for the modeling process described in this paper are not only of importance to academics. According to the Conant/Ashby theorem a management's effectiveness is constrained by the quality of the model on which it is based. Therefore, any managerial activity should be underlaid by a high-quality model. The quality of models could be the ultimate criterion that makes the difference between success and failure in management. The construction of theories is ultimately a natural activity that organizations, like any organism, need in order to be viable. It is the basis for intelligent action and reaction. Excellent managers are conceptual thinkers in this sense; they rely on model-based theory-building for the benefit of their organizations. This leads to a paradoxical conclusion: *Managers must become better theoreticians in order to be superior practitioners.*

ACKNOWLEDGEMENTS

Special thanks go to Rogelio Oliva and three anonymous reviewers for helpful comments.

REFERENCES

Ashby WR. 1956. *An Introduction to Cybernetics.* Chapman and Hall: London.

Barlas Y. 1996. Formal Aspects of Model Validity and Validation in System Dynamics.*System Dynamics Review* **12**(3) : 183-210.

Bell JA, Senge PM. 1980. Methods for Enhancing Refutability in System Dynamics Modeling. In *System Dynamics. TIMS Studies in the Management Sciences*, *14,* Legasto AA, Forrester JW, Lyneis JD (eds). North Holland: Amsterdam; 61-73.

Checkland P. 1981. *Systems Thinking, Systems Practice.* Wiley: Chichester.

Conant RC, Ashby WR. 1981. Every Good Regulator of a System Must Be a Model of that System. In *Mechanisms of Intelligence. Ashby's Writings on Cybernetics,* Conant R (ed). Intersystems Publications: Seaside, CA; 205-214.

Cowles M. 2001. *Statistics in Psychology: An Historical Perspective.* Erlbaum: Mahwah, NJ.

Davis JP, Eisenhardt KM, Bingham CB. 2007. Developing Theory through Simulation Methods. *Academy of Management Review* **32**(2) : 480-499.

Forrester JW. 1961. *Industrial Dynamics.* MIT Press: Cambridge, MA.

Forrester JW. 1968. Industrial Dynamics - A Response to Ansoff and Slevin. *Management Science* **14**(9): 601-618.

Forrester JW, Senge PM. 1980. Tests for Building Confidence in System Dynamics Models. In *System Dynamics,* Legasto AA Jr., Forrester JW, Lyneis JM (eds). North-Holland: Amsterdam; 209-228.

Grove WM, Meehl PE. 1996. Comparative Efficiency of Informal (Subjective, Impressionistic) and Formal (Mechanical, Algorithmic) Prediction Procedures: The Clinical-Statistical Controversy. *Psychology, Public Policy, and Law* **2**(2) : 293-323.

Hannan MT, Pólos L, Carroll GR. 2007. *Logics of Organization Theory. Audiences, Codes, and Ecologies.* Princeton University Press: Princeton and Oxford.

Hanneman RA. 1988. *Computer-assisted Theory Building. Modeling Dynamic Social Systems.* Sage Publications: Newsbury Park, CA.

Henard DH, Szymanski DM. 2001. Why Some New Products Are More Successful than others. *Journal of Marketing Research* **38**(3): 362-375.

Holton EF, Lowe JS. 2007. Toward a General Research Process for Using Dubin’s Theory Building Model. *Human Resource Development Review* **6**(3) : 297-320.

Homer J. 1996. Why we Iterate: Scientific Modeling in Theory and Practice. *System Dynamics Review* **12**(1) : 1-19.

Karlöf B, Lövingsson FH. 2006. *ReOrganisation.* Springer: Berlin.

Lane, DC. 1998. Can We Have Confidence in Generic Structures? *Journal of the Operational Research Society* **49** : 936-947.

Lane DC. 2000. Should System Dynamics be Described as a `Hard' or `Deterministic' Systems Approach? *Systems Research and Behavioral Science* **17** : 3–22.

Lane, DC. 2008. Formal Theory Building for the Avalanche Game: Explaining counter-intuitive behaviour of a complex system using geometrical and human behaviour-al/physiological effects. *Systems Research and Behavioral Science* **25** : in press.

Lane DC, Smart C. 1996. Reinterpreting 'Generic Structure': Evolution, Application and Limitations of a Concept. *System Dynamics Review* **12**(2) : 87-120.

Merton RK. 1957. *Social Theory and Social Structure.* Free Press: New York.

Merton RK. 1968. *Social Theory and Social Structure. Enlarged Edition,* Free Press: New York.

Miller JHP, Page S. 2007. *Complex Adaptive Systems: An Introduction to Computational Models of Social Life*. Princeton University Press: Princeton, NJ.

Neisser U. 1976. *Cognition and Reality. Principles and Implications of Cognitive Psychology.* Freeman: San Francisco, CA.

Paich M. 1985. Generic Structures. *System Dynamics Review* **1**(1) : 126-132.

Patterson CH. 1986. *Theories of Counseling and Psychotherapy,* 4th ed. Harper and Row: New York.

Popper KR. 2002. *The Logic of Scientific Discovery.* Routledge: London.

Randers J. 1980. Guidelines for Model Conceptualization. In *Elements of the System Dynamics Method,* Randers J. Productivity Press: Cambridge, MA; 117-138.

Richardson GP, Pugh AL. 1981. *Introduction to System Dynamics Modeling with DYNAMO.* Productivity Press: Cambridge, MA.

Richardson GP, Andersen DF. 1995. Teamwork in Group Model Building. *System Dynamics Review,* 11(2): 113-137.

Ritchie-Dunham JL. 2002. *Balanced Scorecards, Mental Models, and Organizational Performance: A Simulation Experiment,* PhD dissertation, University of Texas at Austin.

Rogers EM. 2003. *Diffusion of Innovation*. Fifth edition. Free Press: New York.

Schwaninger M. 2003. Modeling with Archetypes: An Effective Approach to Dealing with Complexity. In *Computer Aided Systems Theory - EUROCAST 2003,* Moreno-Díaz R, Pichler F (eds). Springer: Berlin; 127-138.

Schwaninger M, Groesser, SN. 2008. Validation of System Dynamics Models. Encyclopedia of Complexity and System Science. Springer: Berlin (in press).

Schwaninger M, Hamann T. 2005. Theory-Building with System Dynamics. In *Computer Aided Systems Theory - EUROCAST 2005,* Moreno-Díaz R, Pichler F, Quesada-Arencibia A (eds). Springer: Berlin; 56-62.

Senge PM. 2006. *The Fifth Discipline. The Art and Practice of the Learning Organization.* New edition. New York: Doubleday.

Skinner Q (ed). 1985. *The Return of Grand Theory in the Human Sciences.* Cambridge University Press: Cambridge, MA.

Sterman JD. 1984. Appropriate Summary Statistics for Evaluating the Historical Fit of System Dynamics Models. *Dynamica* **10** (2) : 51-66.

Sterman JD. 2000. *Business Dynamics. Systems Thinking and Modeling for a Complex World.* Irwin McGraw-Hill: Boston, MA.

Stouffer SA, Guttman L, et al. 1950. *Measurement and Prediction.* Wiley: New York.

Vennix JAM. 1996. *Group Model Building: Facilitating Team Learning Using System Dynamics.* Wiley, Chichester.

von Foerster H. 1984. *Observing Systems,* 2nd ed. Intersystems Publications: Seaside, CA.

Warren K. Competitive Strategy Dynamics. Wiley, New York.

Wolstenholme EF. 1990. *System Enquiry - A System Dynamics Approach.* Wiley, Chichester.

Wolstenholme EF. 2003. Towards the Definition and Use of a Core Set of Archetypal Structures in System Dynamics. *System Dynamics Review* **19**(1) : 7-26.

Yin RK. 2003. *Case Study Research: Design and Methods*. 3rd Edition. Sage: Beverly Hills, CA.

Zagonel AA. 2004. *Reflecting on Group Model Building Used to Support Welfare Reform in New York State.* Doctoral dissertation SUNY-New York State University: Albany, NY.

Zirger BJ, Maidique MA. 1990. A Model of New Product Development: An Empirical Test. *Management Science* **36**(7): 867-883.

Appendix: Measures to achieve high quality in the modeling and validation process

Quality Criteria	Measures for high quality in the modeling and validation process
1. Refutability	Modeling principles Model formulation as equations Dimensional Consistency Test Parameter Examination Test
2. Importance	Issue Identification Test Adequacy of Methodology Test System Improvement Test System Configuration Test Principles of group model building
3. Precision and clarity	Modeling principles System Configuration Test Model glossary Dimensional Consistency Test Direct Extreme Condition Test
4. Parsimony and simplicity	Modeling principles Boundary Adequacy Structure Test Structure Examination Test Parameter Examination Test
5. Comprehensiveness	Boundary Adequacy Tests Family Member Test Modeling principles
6. Operationality	Principles of group model building Formulation as quantitative model Structure Examination Test Parameter Examination Test Dimensional Consistency Test
7. Validity	Adequacy of Methodology Test System Configuration Test Structure Examination Test Parameter Examination Test Dimensional Consistency Test Extreme Condition Tests Boundary Adequacy Tests Loop Dominance Test Symptom Generation Test Frequency Generation and Phase Relationship Test Behavior Characteristics Test

	Family Member Test Turing Test etc.
8. Reliability	Integration Error Test Family Member Test Methodology Adequacy Test
9. Fruitfulness	Issue Identification Test System Improvement Test Behaviour Sensitivity Test Boundary Adequacy Policy Test Modified Behaviour Test Multiple Modes Test Pattern Anticipation Test Event Anticipation Test Behaviour Anomaly Test Surprise Behaviour Test Family Member Test Principles of group model building
10. Practicality	Modeling Principles Boundary Adequacy Policy Test Symptom Generation Test Sensitivity Tests Principles of group model building